

УДК 004.934.1

Гнида В. – ст. гр. КАм-51

*Тернопільський національний технічний університет імені Івана Пулюя*

## **ПРОБЛЕМИ РОЗПІЗНАВАННЯ МОВИ ЛЮДИНИ**

Науковий керівник: к.т.н. доцент Бадищук В. І.

**Розпізнавання мови** — процес перетворення мовного сигналу в текстовий потік. Часто використовується у наборі технологій, що дозволяють керувати комп'ютером, використовуючи людський голос.

Перший пристрій для розпізнавання мови з'явився в 1952 році, він міг розпізнавати вимовлені людиною цифри. В 1964 році на ярмарку комп'ютерних технологій у Нью-Йорку було представлено пристрій IBM Shoebox.

Збільшення обчислювальних потужностей мобільних пристроїв дозволило і для них створити програми з функцією розпізнавання мови. Серед таких програм варто відзначити Microsoft Voice Command, яка дозволяє працювати з багатьма прикладними програмами за допомогою голосу. Наприклад, можна включити відтворення музики в плеєрі або створити новий документ.

У комп'ютерах Apple Macintosh у налаштуваннях системи є вбудована функція Speech, що здатна аналізувати команди користувача при натисканні певної клавіші, або якщо команді користувача передує ключове слово.

Інтелектуальні мовні прикладні програми, що дозволяють автоматично синтезувати і розпізнавати людську мову, є наступним етапом розвитку інтерактивних голосових систем IVR. Використання інтерактивного телефонного програмного забезпечення в даний час є не даниною моді, а життєвою необхідністю. Зниження навантаження на операторів контакт-центрів і секретарів, скорочення витрат на оплату праці і підвищення продуктивності систем обслуговування — ось тільки деякі переваги, що доводять доцільність подібних програм.

Однак, прогрес не стоїть на місці і останнім часом у телефонних інтерактивних програмах все частіше використовують системи автоматичного розпізнавання і синтезу мови. У цьому випадку спілкування з голосовим порталом стає природнішим, оскільки вибір в ньому може бути здійснений не тільки за допомогою тонового набору, але і за допомогою голосових команд. При цьому системи розпізнавання є незалежними від дикторів, тобто розпізнають голос будь-якої людини. Основною перевагою голосових систем є доброзичливість до користувача — він позбавляється від необхідності продиратися крізь складні і заплутані лабіринти голосових меню. Тепер достатньо вимовлення мети дзвінка, після чого голосова система автоматично перемістить абонента в потрібний пункт меню.

Наступним кроком технологій розпізнавання мови можна вважати розвиток так званих Silent Speech Interfaces (SSI) (Інтерфейсів Безмовного Доступу). Ці системи обробки мовлення базуються на одержанні й обробці мовних сигналів на ранній стадії артикуляції. Даний етап розвитку розпізнавання мови викликаний двома істотними вадами сучасних систем розпізнавання: надмірна чутливість до шумів, а також необхідність чіткої і ясної вимови при звертанні до системи розпізнавання. Підхід, заснований на SSI, полягає в тому, щоби використовувати нові сенсори, що не піддаються впливу шумів, як доповнення до оброблених акустичних сигналів.

Управління об'єктами за допомогою мови відкрило б широкі перспективи перед автоматизацією у багатьох галузях людської діяльності і надало б можливість спілкування з машинами.